

Roll No

**CE-803 (B)/EE-803 (D)/EX-803 (D)/
ME-803(A) (GS)**

B.Tech. VIII Semester

Examination, May 2023

Grading System (GS)

Data Analytics

Time : Three Hours

Maximum Marks : 70

Note: i) Answer any five questions.

किन्हीं पाँच प्रश्नों को हल कीजिए।

ii) All questions carry equal marks.

सभी प्रश्नों के समान अंक हैं।

iii) In case of any doubt or dispute the English version question should be treated as final.

किसी भी प्रकार के संदेह अथवा विवाद की स्थिति में अंग्रेजी भाषा के प्रश्न को अंतिम माना जायेगा।

1. a) What is Big Data? Explain the characteristics of Big Data? 7

बिग डाटा क्या है? बिग डाटा के विशेषताएँ की व्याख्या कीजिए।

b) Differentiate between Regression and ANOVA. 7

प्रतिगमन और अनोवा के बीच अंतर दें।

2. a) Explain with an example how Big Data analytics can be useful in development of smart cities? 7

स्मार्ट शहरों के विकास में बिग डाटा एनालिटिक्स कैसे उपयोगी हो सकता है, इसके बारे में बताइए।

b) Write down the goals of HDFS. 7

HDFS लक्ष्यों के बारे में लिखें।

3. a) Define term frequency and inverse document frequency. 7

टर्म फ्रीक्वेंसी और व्युत्क्रम दस्तावेज़ फ्रीक्वेंसी को परिभाषित करें।

b) Explain the 4 V's of Big Data. 7

बिग डाटा के 4 V's के बारे में व्याख्या करें।

4. a) Explain the architecture and features of Hive. 7

राजमार्ग की वास्तुकला और विशेषताओं की व्याख्या कीजिए।

b) Explain the concept of metastore in Hive. 7

हाईवे में मेटास्टोर की अवधारणा को स्पष्ट कीजिए।

5. a) What is Pig Latin and why do we need it? 7

Pig Latin क्या है? और हमें इसकी आवश्यकता क्यों है?

b) Explain Map Reduce programming model. 7

मैप रिड्यूस प्रोग्रामिंग मॉडल को समझाइए।

6. a) Explain Hadoop Architecture with diagram and components. 7

हडूप आर्किटेक्चर को डायग्राम और कंपोनेंट्स के साथ समझाइए।

b) What is Zoo keeper? Write the advantages of Zoo Keeper. 7

Zoo keeper क्या है? Zoo keeper के लाभ के बारे में लिखिए।

7. a) Explain any three Hive QL DDL commands with syntax and example. 7

तीन हाईवे QL DDL कमांड को सिंटैक्स और उदाहरण के साथ समझाइए।

- b) Explain the Indexing and Aggregation for Mongo DB. 7
मोंगो DB के लिए अनुक्रमण और एकत्रीकरण की व्याख्या करें।

8. Write short notes on any three of the following: 14

- a) Capacity scheduler in Map Reduce
- b) Information Management
- c) Metastore in Hive
- d) Term frequency

निम्नलिखित में से तीन पर संक्षिप्त नोट्स लिखिए।

- अ) Map Reduce में क्षमता निर्धारित
- ब) सूचना प्रबंधन
- स) Hive मेटास्टोर
- द) टर्म फ्रीक्वेंसी
